Cooley Godward LLP
Attorney Docket No.: CNTW-004/00US
Client No.: C6958-2005

APPLICATION FOR PATENT

TITLE: **DYNAMIC CONFIGURATION OF NETWORK DEVICES TO ENABLE DATA TRANSFERS**

INVENTOR(S): **GLEN D. TINDAL AND JEFFERY A. SCHENK**

CERTIFICATE OF MAILING

I hereby certify that this correspondence or fee is being deposited with the United States Postal Service via "Express Mail", Express Mail No.: EL621462614US to addressed to: Box Patent Application, Assistant Commissioner for Patents, Washington, D.C. 20231 on this date: December 6, 2000

By: _____

101888 v1/BD
26M801! DOC
120600

RELATED APPLICATIONS

[001] The following commonly owned and assigned patent applications are hereby incorporated by reference in their entirety:

1) Patent Application No. CNTW-001/00US, entitled ***System and Method for Configuration, Management and Monitoring of Network Resources***, filed on December 6, 2000:

2) Patent Application No. CNTW-002/00US, entitled ***System and Method for Redirecting Data Generated by Network Devices***, filed on December 6, 2000;

3) Patent Application No. CNTW-003/00US, entitled ***Event Manager for Network Operating System***, filed on December 6, 2000;

4) Patent Application No. CNTW-005/00US, entitled ***Network Component Configuration and Management Method***, filed on December 6, 2000; and

5) Patent Application No. CNTW-006/00US, entitled ***Network Operating System Data Directory***, filed on December 6, 2000.

FIELD OF THE INVENTION

[002] The present invention relates generally to network systems. More particularly, but not by way of limitation, the present invention relates to systems and methods for dynamic configuration of network devices to thereby enable efficient data transfers.

BACKGROUND OF THE INVENTION

[003] Data and the effective, timely movement of data has become the lifeblood of many modern enterprises. Unfortunately, network infrastructure limitations are impinging upon the ability of enterprises to timely move data. Enterprises that require data to be delivered within very strict time requirements are being most severely impacted by these network infrastructure limitations. To guarantee the timely delivery of their data, these enterprise often are forced to pay steep prices. Moreover, network providers are being forced continually to upgrade their infrastructure to supply even the basic services to their customers. Accordingly, the networking community is searching for a method and system to better and more economically utilize the existing network infrastructure, thereby improving the transfer of data and reducing the associated cost.

[004] With regard to the actual transmission of data, enterprises are searching for a way to pay only for the bandwidth that they use. In essence, they are looking to optimize the use of bandwidth. Presently, an enterprise that requires the bandwidth provided, for example, by a T1 line may be forced to rent a dedicated T1 line for an entire month even

though the enterprise may only need the T1 line for a few days within that month. Renting the T1 line for such an extended period is wasteful on two fronts. First, the enterprise is paying for many days of service that it does not use. Preferably, an enterprise should only pay for the service that it actually uses. Second, the bandwidth available on the T1 line for those unused days is wasted because other enterprises do not have access to it. If this unused bandwidth can be captured and made available to other enterprises, the existing network infrastructure can be better utilized to meet the demands of more enterprises. Unfortunately, no device or method exists to effectively optimize the provisioning of bandwidth. Thus, even though dedicated lines are expensive and cumbersome, companies requiring rapid, predictable transfer of data presently have no other acceptable option.

[005] With regard to the actual routing of data, enterprises are searching for an efficient way to route data based upon priority. Several methods have been developed to aid in routing data based upon priority. These present methods, however, are not completely satisfactory. For example, a feature called weighted fair queuing can be enabled on modern routers. This feature requires that the router read a precedence bit from each packet being passed through the router and then queue lower priority packets while routing higher priority packets. Although weighted fair queuing can be effective, it requires a great deal of processing power and memory within each router, and when enabled for a long period of time, weighted fair queuing can cause a router to crash. Additionally, because weighted fair queuing requires a router to analyze each packet that it receives, it slows the overall operation of the router and, thus, the network.

[006] Because router configuration is a somewhat complicated and time consuming process, especially if numerous routers are involved, network administrators tend to configure routers in the network either to use weighted fair queuing at all times or not to use weighted fair queuing at all. In reality, most network administrators would like to use weighted fair queuing some of the time and disable it the rest of the time. Unfortunately, selectively enabling and disabling weighted fair queuing is so cumbersome that it cannot be effectively implemented. Accordingly, a device and method are needed to selectively enable router optimization techniques such as weighted fair queuing.

[007] Because the present network technology suffers from significant drawbacks, a solution is needed that can efficiently and effectively optimize a network to enable a more efficient transfer of data. In particular, a system and method are needed in which both the bandwidth usage and/or the router performance can be easily optimized. Such a system and method would not only address the needs of the network community, but also provide new advantages such as content transfer optimization.

SUMMARY OF THE INVENTION

[008] To address the problems and limitations of present network technology, the present invention provides for an efficient, effective optimization of a network to enable data transfers. In particular, but not by way of limitation, the present invention provides a

method and apparatus to optimize bandwidth usage, routing performance and content delivery.

[009] In one embodiment, for example, a network provider (or manager) can receive a request to transfer a block of data between two points. Such a request could indicate the identity of the party requesting the transfer and the volume of data to be transferred. The network provider could then identify the services to which the requesting party is entitled. For example, the network provider could determine whether the data block should be transferred with a high priority, a medium priority, or a low priority.

[010] After the network provider has determined the appropriate level of service to assign to the data transfer, it can identify the path and associated network devices for transferring the data. Next, using a system in accordance with the present invention, the network devices along that path can be dynamically configured to handle the data transfer within the appropriate service level. For example, priority data handling features (such as weighted fair queuing) can be enabled on the appropriate routers. Additionally, or even alternatively, a virtual, dedicated line between the two transfer points can be established by reconfiguring the appropriate optical devices. Once the data transfer has been completed, the priority data handling features can be disabled and/or the virtual dedicated line between the two points can be torn down.

[011] Accordingly, in the above-described embodiment, network resources can be provisioned "just-in-time." Moreover, when network resources are not being used, they

can be returned to the pool of available resources, and when special data handling features such as weighted fair queuing are not needed, they can be turned off.

BRIEF DESCRIPTION OF THE DRAWINGS

[012] Various objects and advantages and a more complete understanding of the present invention are apparent and more readily appreciated by reference to the following Detailed Description and to the appended claims when taken in conjunction with the accompanying Drawings wherein:

[013] FIGURE 1 illustrates a present network system connecting portions of an enterprise with a dedicated line;

[014] FIGURE 2 illustrates a dynamically configurable network system, in accordance with the present invention, that can be optimized for efficient data transfers;

[015] FIGURE 3 is a flowchart of the process of bandwidth optimization on a network system such as the one in FIGURE 2;

[016] FIGURE 4 is a flowchart of the process of routing optimization on a network system such as the one shown in FIGURE 2; and

[017] FIGURE 5 is a flowchart of the process of content transfer optimization on a network system such as the one shown in FIGURE 2.

DETAILED DESCRIPTION

[018] Although the present invention is open to various modifications and alternate constructions, a preferred exemplary embodiment that is shown in the drawings is described herein in detail. It is to be understood, however, that there is no intention to limit the invention to the particular forms disclosed. One skilled in the art can recognize that there are numerous modifications, equivalents and alternative constructions that fall within the spirit and scope of the invention as expressed in the claims.

[019] Referring now to FIGURE 1, there is illustrated a present network system 100 connecting two portions of an enterprise (105a and 105b) with a statically defined, dedicated line 110. As previously discussed, the enterprise may be forced to rent the dedicated line 110 for an entire month even if the line 110 is only used for a few days of that month. Thus, the line 110 (or at least the provisioned portion of the line 110) can sit idle for the majority of the time. Obviously, by allowing the line 110 to sit idle, valuable network resources that other enterprises could utilize are wasted, and the enterprise renting the line is forced to pay for services that it is not using.

[020] Referring now to FIGURE 2, there is illustrated a dynamically configurable network 115 that can be optimized for efficient data transfers. In this embodiment, the two portions of the enterprise are connected to a network 115 that includes a plurality of routers 120 and optical devices 125. (As one skilled in the art can understand, the configuration of the routers 120 and optical device 125 in FIGURE 2 is merely

exemplary.) Rather than renting a statically defined, dedicated line 110 (shown in FIGURE 1), the enterprise can request that a virtual, dedicated line be temporarily provisioned within the network 115. This concept of requesting and providing a virtual, dedicated line on demand may be referred to as "just-in-time provisioning." For example, an enterprise could request, from the network provider, a guaranteed bandwidth of 1.544 Mbps (equivalent to a T1 line) between two points for a period of two days starting in five minutes. Normally, the network provider could not fill such a request because configuring the network to provide such a bandwidth would take significantly longer than five minutes. Moreover, a network provider could not fill such an order because, with present technology, it would not be economically feasible to establish the requested service for such a short period of time. In fact, configuring a path to reserve that amount of bandwidth can take weeks with the present technology.

[021] Using the present invention, however, the network provider could perform just-in-time provisioning and provide enterprises with the requested bandwidth for the requested time frame. Network providers can perform this just-in-time provisioning through a dynamic configuration of the relevant network devices. Assume, for example, that optical device 125b and its associated lines could provide the bandwidth requested by the enterprise. This optical device could be dynamically identified and dynamically configured to reserve the requested bandwidth for the requested timeframe. Moreover, router 120a and router 120f could be dynamically configured to route data from the enterprise to optical device 125b rather than to any other network device and associated path.

[022] In one embodiment of the present invention, the dynamic configuration of network devices is achieved through directory-based networking. One example of directory-based networking is described in commonly owned and assigned patent application no. CNTW-001/00US, entitled *System and Method for Configuration, Management and Monitoring of Network Resources*, filed on December 6, 2000. Briefly, directory-based networking involves storing a configuration record for each network device in a central repository. When a network device needs to be reconfigured, the centrally-stored configuration record for that device can be retrieved and altered. The altered configuration record can then be used to generate the device-specific code needed to reconfigure the relevant network devices. Finally, once the device-specific code has been generated, that code is provided (either through a push or get) to the appropriate network device(s). Thus, by using directory-based networking, network devices can be dynamically configured with a minimal amount of actual human intervention, thereby allowing for just-in-time provisioning of network resources.

[023] Referring now to FIGURE 3, there is illustrated a flowchart of the process of bandwidth optimization on a network system such as the one in FIGURE 2. In particular, FIGURE 3 describes the process for optimizing the utilization of a fiber optic line or any other type of line. Initially, an enterprise requests a bandwidth of a certain size between two points for a particular timeframe (or for a particular volume of data). Next, a path 305 and the associated network devices that can provide the requested bandwidth are identified 310. The network devices along that path are then configured to provide the requested bandwidth 315. For example, the optical devices along the identified path can

be configured to reserve the requested bandwidth for the requesting enterprise. Once the requested timeframe has expired, the optical devices can be returned to a default setting, thereby tearing down the temporary dedicated path 320. The network resources previously dedicated to the enterprise's path are now returned to the pool of network resources where they can be accessed by other enterprises.

[024] Accordingly, one embodiment of the present invention provides a method for just-in-time provisioning of network resources. In particular, this embodiment provides a method for easily and dynamically establishing and tearing down a virtual, dedicated transmission path. With the present invention, enterprises can request and pay for only those network services that they need. Moreover, through the present invention, network providers can better utilize network resources by returning unused network resources to a pool of generally available resources. In prior art systems, these unused network resources could have remained dedicated to a single enterprise whether or not they were actually being used.

[025] Referring now to FIGURE 4, there is illustrated a flowchart of the process of routing optimization on a network system such as the one shown in FIGURE 2. In this embodiment, routers and similar devices can be optimized to efficiently handle data based upon the priority of the data. As with bandwidth optimization, routing optimization can be enabled by directory-based networking principles that allow for dynamic configuration of network devices. In this particular embodiment, prioritization

features of routing devices can be selectively enabled and disabled to route data based upon priority of that data.

[026] Initially, an enterprise may request that data be routed with a high priority or that the enterprise be given a certain routing priority for a particular timeframe 405. In either case, the network provider can determine the enterprise's service level and link that service level to a priority 410. For example, if the enterprise has a top level service agreement, that enterprise may have access to the highest level of priority that the network provider can give. Alternatively, if the enterprise has a lower level service agreement, that enterprise may only have access to a mid-level priority. By tiering priority in this fashion, enterprises can select and pay for the level of service that they need. Moreover, network providers can maximize network resources and revenue by providing higher priority service(s) to those customers that need it and that are willing to pay for it.

[027] Responsive to an enterprise requesting that data be routed with a certain priority, a pathway can be identified 415 and routers along that pathway can be dynamically reconfigured to enable priority data handling features such as weighted fair queuing 420. With these priority data handling features enabled, higher priority data can be routed before lower priority data.

[028] As previously described, weighted fair queuing is effective for routing data based upon priority but can cause network congestion and router failure when used

unnecessarily. The present invention addresses this problem by allowing for weighted fair queuing and similar priority data handling features to be dynamically disabled 425. In other words, weighted fair queuing can be turned off when not needed without significant difficulty.

[029] Notably, the present invention allows for the concurrent operation of bandwidth optimization and routing optimization. For example, an enterprise could request a virtual dedicated line and also request routing priority to that line. The network provider could provision these services based upon a service level agreement with the enterprise.

[030] Referring now to FIGURE 5, there is illustrated a flowchart of the process of content transfer optimization on a network system such as the one shown in FIGURE 2. In this method, an enterprise can notify its network provider that it has a block of data to be moved between two points within certain quality parameters 505. The quality parameters can be directly indicated, e.g., specified delivery time, or they can be determined according to the type of content being transferred 510. For example, movement of disk mirroring content may take a higher priority than movement of replication content.

[031] After the request for content transfer has been made and the importance of that transfer has been determined 515, the path for transferring that content can be identified 520. If the requesting enterprise is utilizing a virtual dedicated line, the content transfer can be made using that line. Additionally, the routers connected to that line can be

configured such that weighted fair queuing is enabled 525 for this content transfer and disabled when the transfer is complete 530. Alternate embodiments involve differing combinations of bandwidth optimization and content optimization. As with bandwidth optimization and routing optimization, content-delivery optimization can be achieved, in a variety of ways, including through the use of directory-enabled networking.

[032] In conclusion, the present system provides, among other things, a system and method for optimizing the utilization of network resources. Those skilled in the art, however, can readily recognize that numerous variations and substitutions may be made in the invention, its use and its configuration to achieve substantially the same results as achieved by the embodiments described herein. Accordingly, there is no intention to limit the invention to the disclosed exemplary forms. May variations, modifications and alternative constructions fall within the scope and spirit of the disclosed invention as expressed in the claims.